Current Biology, Volume *24*

Supplemental Information

Human-Mediated Loss of Phylogenetic and Functional Diversity in Coral Reef Fishes

Stéphanie D'agata, David Mouillot, Michel Kulbicki, Serge Andréfouët, David R. Bellwood, Joshua E. Cinner, Peter F. Cowman, Mecki Kronen, Silvia Pinca, and Laurent Vigliola

Supplemental Data

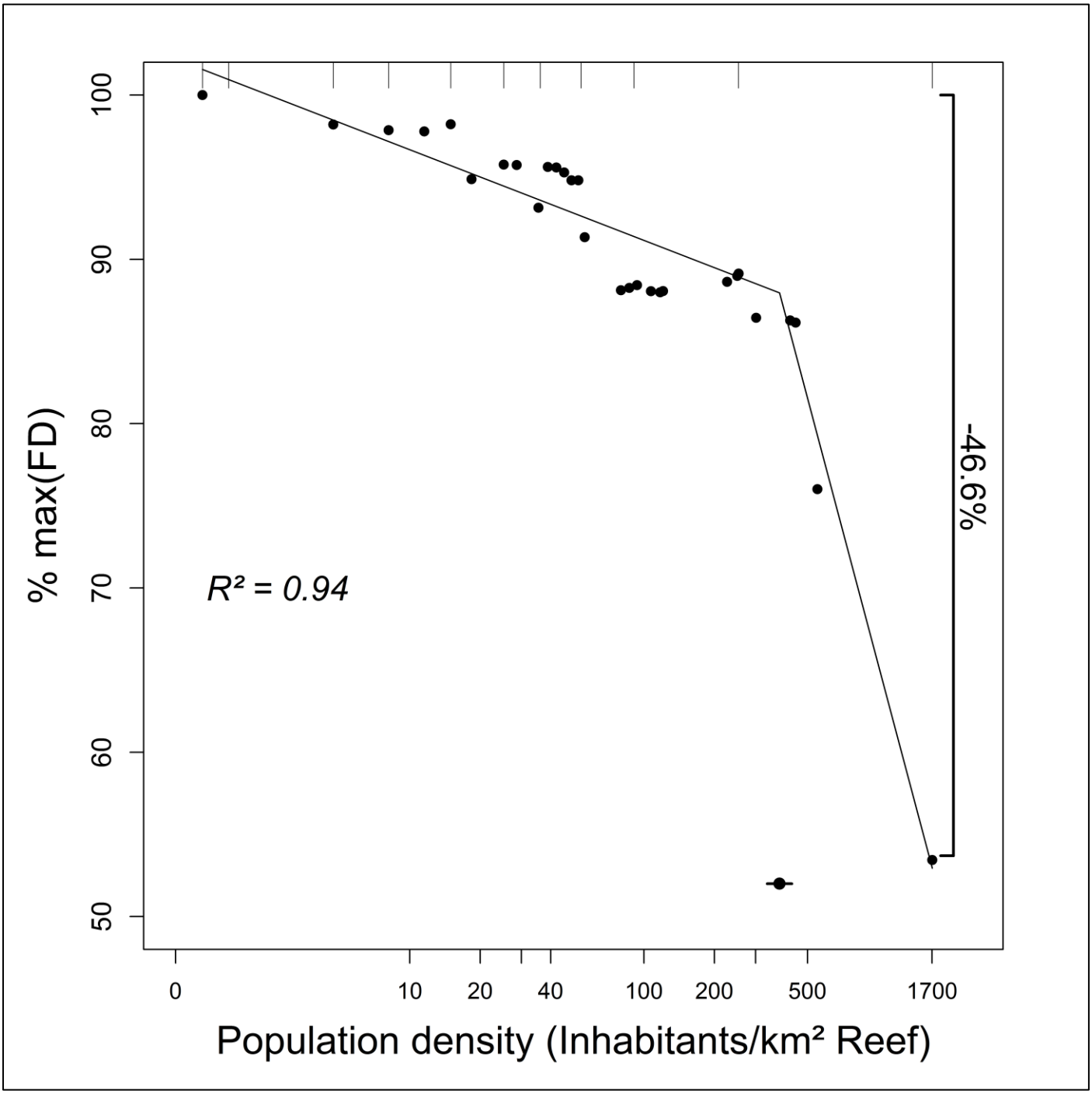

Figure S1 [Related to Figure 3]: Partial dependence plot of FD for parrotfish along population density per km² of coral reef.
Fitted variations of FD of parrotfish for the population density per km² of coral reef (X-axis), one of the most influential factor that predicts FD of using BRT model. Fitted variations were predicted using Biogeography and Habitat PCoA axes and Population density as predictors in the BRT model (Figure 3 d,e,f). Values were predicted for each factor, holding values for all other factors at their mean value for the human density.
Black lines represent the alternative fitted model to Figure 3, considering all sites.
Breaking-point estimates (380.7) and 95% confidence intervals (337.3 - 429.2) are plotted. Davies' test for a change of slope: 380.9 (***) (significant at p-value < 0.001).

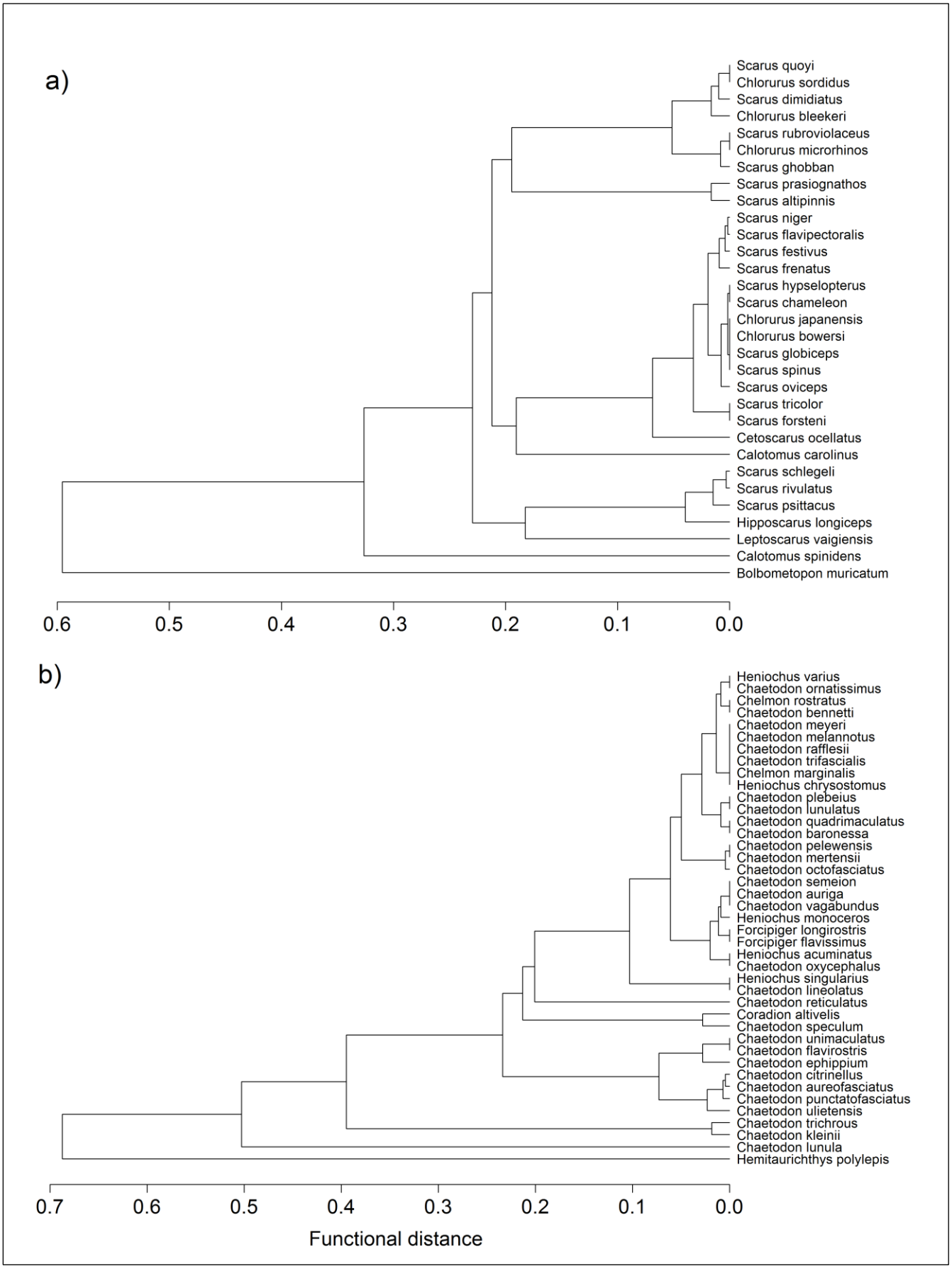

Figure S2 [Related to Figure 2&3]: Functional tree of Parrotfish (31 species) (a) and Butterflyfish (41 species) (b).

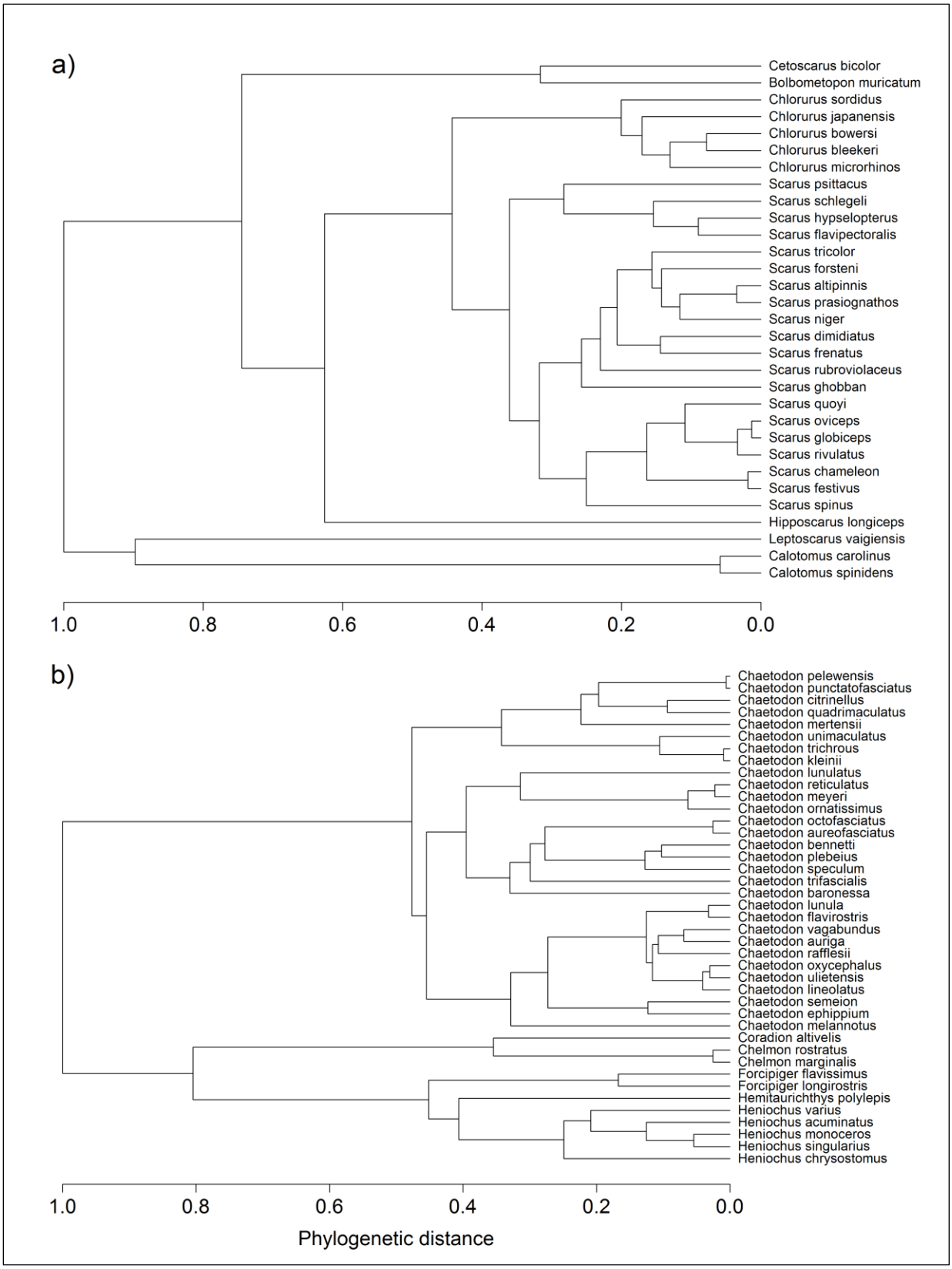

Figure S3 [Related to Figure 2&3]: Phylogenetic tree of Parrotfish (31 species) (a) and Butterflyfish (41 species) (b).

Table S1 [Related to Figure 2]: Parameters and performance of i) simplified BRT models for Parrotfish **(a)** and Butterflyfish **(b)** and ii) simplified BRT models using population density as the only Human variable for Parrotfish **(c)** for Species Richness (S), Phylogenetic Diversity (PD) and Functional Diversity (FD). n is the number of PCoA axes used as explanatory factors after the simplification procedure (see Supplemental Experimental Procedures), lr is the learning rate, tc is the tree complexity, N.trees is the number of optimal trees. CV R^2 is the correlation coefficient from the cross-validation procedure with n folds (nf) equal to 20 and bag fraction (bf). SE CV R^2 is the standard error.

a)	n	lr	tc	N.trees	CV R^2	SE CV R^2	bf	nf
S	6	0.002	10	1650	0.58	0.019	0.7	20
PD	7	0.002	11	2900	0.51	0.019	0.7	20
FD	6	0.004	9	2000	0.54	0.022	0.7	20

b)	n	lr	tc	N.trees	CV R^2	SE CV R^2	bf	nf
S	6	0.004	8	2300	0.62	0.014	0.7	20
PD	5	0.002	11	2900	0.50	0.027	0.7	20
FD	5	0.002	11	1250	0.39	0.014	0.7	20

c)	n	lr	tc	N.trees	CV R^2	SE CV R^2	bf	nf
S	6	0.002	10	3150	0.58	0.017	0.7	20
PD	7	0.002	11	2050	0.49	0.020	0.7	20
FD	6	0.004	9	1600	0.53	0.021	0.7	20

Table S2 [Related to Figure 2]: Mean, Standard deviation (Sd) and range of S, PD and FD for Parrotfish and Butterflyfish.

	Parrotfish			Butterflyfish		
	Mean	Sd	Range	Mean	SD	Range
S	6	2.95	2 - 17	6	2.83	2 – 20
PD	0.74	0.28	0.0042 - 1.3386	0.83	0.38	0.01 - 1.41
FD	0.38	0.15	0.00035 - 0.66989	1.3	0.1	0.69 - 0.77

Table S4 [Related to Figure 2]: Life history traits used to construct functional trees [S6].

Trait	Values	Functional significance
Maximum size	Centimeters (cm)	Metabolic rate, Home range, ingestion production, secondary production [S10]
Diet-class	C: Invertebrates without size distinction C1: Macro-invertebrates C2: Macro-invertebrates and sessile invertebrates (corals, sponges, etc.) H: plant material without distinction (algae, phanerogam, etc.) H1: Macro algae H2: Micro-algae, turf, cyanobacteria D: Detritus, inert organic material P: Nekton (fish, living cephalopods) Z: Plankton	Diet composition, susceptibility to predation [S10]
Home-Range	S: Sedentary M: Mobile W: Wide Range	Feeding type, foraging method, response to predation risk or prey defense [S10]
Activity	N: Nocturnal D: Diurnal B: Both	Susceptibility to predation [S10]
Schooling	S: Solitary P: paired F: Small School M: Medium School L : Large School	Susceptibility to predation [S10]
Level	B: Bottom L: Low above bottom H: High above bottom	Feeding type, foraging method [S10]

Supplemental Experimental Procedures

Methods

Sampling

Reef fish fauna and associated coral reef habitats were surveyed from 2002 to 2009 in 63 sites across the Pacific among 17 territories (Figure 1), spanning ~35° latitude and ~85° longitude. Data were collected along 1553 underwater visual transects, all located in the fishing ground of villages (sites).

The sampling design consisted in 4-6 sites per country and 24 transects per site, with a stratified design among the main geomorphologic structures or reef types present at a given site: 1) sheltered coastal reef, 2) intermediate reef, 3) back-reef (inner/lagoon side of barrier reef) and 4) outer reef (ocean side of fringing or barrier reefs). For each geomorphologic structure, transects were performed both on the reef flat (3) and slope (3) when feasible. Transects were performed parallel to the reef between 0 to 15 meters depth. Maps from the Millennium Coral Reef Mapping Project satellite images [S1], and in situ observation at dive sites allowed identification of these four geomorphologic structures and calculation of reef areas at each site.

Distance-sampling Underwater Visual Census (D-UVC) [S2] technique was used to survey finfish along 50 meters long transects in selected sites. Briefly, the technique consists in recording fish size and abundance at the species level and the perpendicular distance of each fish or group of fish to the transect line. D-UVC datasets were then truncated at a distance of 5 meters on each side of transects allowing the calculation of abundances, biomasses and diversities over 500 m² transects (2 sides x 5 m width x 50 m long) [S2].

This study focuses on 31 species of Parrotfish (Scarinae) and 41 species of Butterflyfish (Chaetodontidae) that were selected for their documented responsiveness to habitat quality [S3] and human impact (e.g. [S4, 5]). Species lists for each group are listed in Figure S2&S3.

Biodiversity components

Species richness is the number of species per 50x10 meters transects (number of species per 500m²). Biomass (g/m²) was calculated by multiplying fish abundance on each transect by individual weight ($W = aL^b$) with W the weight in g, L the individual fork length (cm), a and b being the weight/length parameters [S6] .

A functional diversity metric based on species diet and feeding behavior was selected. To this aim we retained six functional traits linked to these two aspects: 1) Maximum size, 2) Diet, 3) Home-Range, 4) Position over the reef, 5) Activity, 6) Gregariousness. The details and functional role of each trait are listed in Table S4. The functional traits database has benefited from several decades of observations in the Indo-Pacific ([S6–9], completed by FishBase[1] for 8/31 species for parrotfish and 8/41 species for butterflyfish.

The functional trees (Figure S2) for each species group were built using the six functional traits. The functional traits are both quantitative and qualitative and the Gower distance was used to estimate distances between species pairs in functional trees, one for each group. Because the clustering procedure has an influence on functional diversity estimate [S11], we carried out an optimization procedure to select the tree that minimizes the dissimilarity between the initial functional distance matrix between species pairs and the cophenetic distance matrix calculated from the tree [S11].

The phylogenetic trees for parrotfish and butterflyfish (Figure S3) are from [S12]. Briefly, a range of published and unpublished sequence data for the parrotfish and butterflyfish were obtained from GenBank. Gene datasets were then constructed for each group of species and were concatenated into supermatrices for standard phylogenetic analysis [S12].

[1] fishbase.org

Biomass Weighted Phylogenetic Diversity (PD) and Functional Diversity (FD) were computed based on the phylogenetic entropy index [S13].

Let A_{ik} be the abundance (number of individuals or biomass) of species *i* within transect *k*, the regional pool having a total of *S* species and *K* transects. The weight of each transect *k* is proportional to its relative abundance over all transects:

$$f_{.k} = \frac{\sum_{i=1}^{S} A_{ik}}{\sum_{k=1}^{K}\sum_{i=1}^{S} A_{ik}} \qquad \text{(Eq. 1)}$$

Phylogenetic diversity or "local phylogenetic entropy (within transect *k*)", following [S13] is given by:

$$PD_k = H_k = -\sum_{t=1}^{T} l(b_t) \times p_k(b_t) \times \ln\left[p_k(b_t)\right] \qquad \text{(Eq. 2)}$$

where *T* is a phylogenetic tree for the species regional pool, $l(b_t)$ is the length of a branch b_t, and $p_k(b_t)$ is the local proportion of abundance given by:

$$p_k(b_t) = \frac{\sum_{i=1}^{S_t} A_{ik}}{\sum_{i=1}^{S} A_{ik}} \qquad \text{(Eq. 3)}$$

where S_t is the number of species descending from b_t. In the formulae above, phylogenetic diversity was calculated using fish biomass as the measure of species abundance. As such, the PD index used in this study represents a biomass weighted phylogenetic diversity index (PD).

Original variables

Human variables

Socio-economic data collection was performed at each site using a standard set of questionnaires, which included a household survey (key socioeconomic parameters and consumption patterns), finfish fisheries survey, finfish marketing survey and general information survey (for key informants). Socio-economic surveys were conducted between March 2002 and December 2009 by the same team of two scientists with the help of local attachments. The questionnaires were designed to allow a minimum dataset to be developed for each site with: 1) the community's dependency on marine resources to be characterized, 2) assessment of the community's engagement and the possible impact of finfish and invertebrate harvesting and 3) comparison of socioeconomic information with data collected through finfish resource surveys [S21].
The socio-economic dataset was complemented by demographic data available from a global database (collected from the Socioeconomic Data and Applications Center SEDAC http://sedac.ciesin.columbia.edu). Specifically, human population occurring within a 10-km ("Human 10 km"), 20km ("Human 20 km") and 50km ("Human 50 km") radius of each transect retrieved from the demographic database for the year 2000. These radii included the human population outside the village, thus the potential capitals surrounding the village. The human population in a buffer represents a realistic human pressure by adding a spatial gradient of human population from a highly located (the village) to a regional scale.
In PNG, McClanahan and Cinner (2008) observed that overfishing of the higher value and high trophic-level species occur in areas close to selling market. Thus probable selling markets in each country were located for each site from literature and socio-economic interviews and distances between fishing sites to those markets were then retrieved as the crow flies with Google Earth®.
In addition, economic index at the country level were added.

Gross Domestic Product (GDP) is the total value of all goods and services a nation produces in one year (generally in dollars). The GDP per capita (per person) is often used as a measure of standard of living [S22].
Variables are listed in Table S5a).

Biogeographic variables

Biogeographic factors were included in the models to take into account the geographical extent of the study. Sandin et al (2008) [S15] demonstrated the positive effect of island area which describes both habitat availability and habitat diversity. Inversely, isolation has a negative influence on species richness through the decreasing immigration and increasing extinction rates [S15].
Thus the following metrics were included: Island size (m^2), distance to nearest island (km), distance to the nearest group of islands (km) and distance to the nearest continent, retrieved for each site/island (collected from the United Nations Environment Programme (UNEP) Island Directory; http://islands.unep.ch/Tisolat.htm). Island type was characterized by three categories: high island (island without lagoon, which include tectonically uplifted reefs, such as Nauru), low island (island with a large lagoon) and atoll (no island except reef islands which are islands that are created by the accumulation of reef sediments).
Reef area is a measurement of habitat availability and connectivity (e.g. [S15, 16]) and was added in the models. Reef area (km^2) in a buffer radius (km) was computed using the Global Distribution of Coral reefs (2010) (collected from the United Nations Environment Programme (UNEP) World Conservation Monitoring Centre www.unep-wcmc.org).
Seven buffers were retained to take into account the ecological and evolutionary scale of connectivity. We used 10km, 20km, 70km [S16], 150km [S17] and 300km buffers [S18, 19] as proxy of ecological connectivity and 600km [S20] and 1000km as proxies for evolutionary connectivity.
Longitude and latitude were recorded for each transect.
Variables are listed in Table S5b).

Habitat variables

The Medium Scale Approach (MSA) was used to record substrate characteristics along transects where finfish were counted by D-UVC. MSA has been developed by [S14] to specifically complement D-UVC surveys. The method consists in recording depth, habitat complexity, and 23 substrate parameters (e.g. % coral cover) within ten 5x5 m quadrats on each side of 50-m transects, for a total of 20 quadrats per transect. Habitat characteristics of each transect are then calculated by averaging over the 20 quadrats each habitat parameter potentially relevant to explain the structure of finfish communities.
Shannon substrate diversity index for coral cover was computed using the percentages of the different shapes (encrusting, massive, digitate, branching, foliose, tabulate) of live coral cover.
Habitability is a semi-qualitative variable between 1 (low) and 4 (high). It is a compromise between relief, coral cover and the complexity. It summarizes the substrate/habitat quality and is assigned by the diver.
Variables are listed in Table S5c).

Statistical analysis

Linear relationships are usually assumed between explanatory and response factors whereas they are often nonlinear and even non-monotonic in nature, including threshold effects [S23–26]. Additionally, interactions between explanatory factors cannot be fully disentangled using traditional modeling methods. This is a failure considering that confounding effects have been largely observed between environmental factors [S27]. Explanatory power of ecological models is thus largely affected, inducing distorted conclusions of studies and leading to inappropriate management decisions [S23].
As an alternative, regression trees have proven to offer higher predictive performances by incorporating with success thresholds and interaction effects (e.g. [S23]). We used Boosted Regression Trees (BRT) to model biodiversity components. This technique minimizes the deviance between predicted and observed

data by generating at each step a new tree explaining the residuals from the previous and so until the last tree [S28]

We decided, in agreement with Elith et al (2008) [S28] that a low cross-validation (CV) predictive deviance hence a high correlation coefficient from cross-validation (CV R^2) as well as a sufficient number of trees (N trees) (> 1000) were essential criteria to select a good BRTs model.
To reach this goal, a set of key parameters needed to be determined to run BRTs and to generate an optimal model. Firstly, the learning rate (lr) or shrinkage parameter determines the contribution of each tree to the growing model [S28]. Secondly, the tree complexity (tc) determines the depth of the interactions between factors of each tree [S28]. Those two parameters will determine the optimal number of trees needed to fit the model. Since decreasing lr increases the number of trees required, and increasing tc decreases the number of trees necessary, the optimal number of trees and the minimum deviance results from a trade-off between lr and tc. Consequently, a combination of lr (0.005, 0.004, 0.003, 0.002) and tc (1 to 20) were used to run BRT models for S, PD and FD for each group of species. A 20-fold cross-validation (CV) was performed using a bag fraction (bf) from 0.5 to 0.7 (increment 0.1) and predictive performance was determined using the out-of-bag (OOB) estimate. OOB estimates of error rate are based on bootstrap sampling using a random subset of records (bf) as training data for each iteration. A total of 100 models were run for each biodiversity index and for each group of species.
We chose the optimal combinaison of tc, lr and bf as the one minimizing OOB estimates of error rate with a N tree superior to 1000.
Contributions of each factor (%) are the proportion of each factor selected to split the data among all the trees, weighted by the squared improvement to the model as a result of each split, and averaged over all trees [S28]. The factors with the highest percentage for contribution are the most important factors contributing to the model. Elith et al. (2008) [S28] suggests that explanatory factors with a relative influence of less than the 5% threshold usually do not improve the prediction capacity of the model and should not be considered.
Once each model has been run and selected relatively to its predictive deviance, a modified simplification procedure inspired by Elith et al (2008) was applied so irrelevant factors were removed objectively from the model [S28].
Usually, removing the factors with the lowest contributions of BRTs did not modify clearly the deviance until reaching a threshold at which the removal of the next less important factor increased significantly the change in CV predictive deviance.
The number of factors to remove was thus determined by a net increase in the change of predictive deviance, determined thanks to breaking points regression which helped to determine the break point of predictive deviance. This simplification procedure was repeated 20 times for each model and the number of factors to be removed with the highest frequency was selected. Then, a new model was run using the simplified set of factors.

Human (19) (Table S5a), Biogeographic (17) (Table S5b) and Habitat (33) (Table S5c) original variables were available to model S, PD and FD. The number of original variables in each category, however, is heterogeneous potentially leading to an overrepresentation of one category (here Habitat) over another. This would bias the overall contribution of each category, which is based on the summed contribution of each original variables for each category. BRT also require independent explanatory variables.
Thus to determine the main contribution of each category of factors independently, we performed Principal Coordinate Analysis (PCoA) to select 5 axes for each category as explanatory factors.

All BRT were conducted in R (R Development Core Team 2011 version R version 2.14.0) using gbm package version 1.6-3.1 and custom code available online [S28].

Characterization of PCoA axes with original variables.

Because qualitative and quantitative original variables were used, scores of each original variable could not be computed to characterize PCoA axes.
The strength of the relationship between original quantitative variables and PCoA axes, using the Pearson correlation coefficient (R), was used to assign a score to each original quantitative variable for each PCoAs axis (Table S3).
The median coordinates for the different modalities of qualitative original variables (with the first and last quartiles (in brackets) were used to define the original qualitative variables characterizing PCoA axes.
Table S3 summarizes (Table S3).

Procedure to extract the marginal effect of population density on biodiversity components of parrotfish

Population density characterizes Human PC1, which is among the most influential PCoA axis (Table S3) for phylogenetic and functional diversity (Figure 2, Table S3).
To assess the pure effect of population density on all three biodiversity components of parrotfish, we rerun BRTs using 1) the set of Habitat and Biogeographic PCoA axes the most influential for each biodiversity component (Figure 2, Table S4) and 2) replacing the Human PC1 by the population density. Since information on population density can be contained in others Human PCs, no others were included in the models.
The marginal effect of Human density on parrotfish's species richness, phylogenetic and functional diversity was estimated after accounting for the average effects of all other variables in the model [S28].
Table S1c) summarizes the performances of the new models.

A broken-line relationship between human density and each biodiversity component for parrotfish were considered. The suitability of this type of model was tested through the null hypothesis of no change in slope with Davies' test for difference-in-slope [S29] as implemented in the Segmented package version 0.2-9.4 and conducted in R (R Development Core Team 2011 version R version 2.14.0.).
For the biodiversity component where significant variation in slope was detected, a broken-line model was applied using the segmented function in the segmented package [S30].

Supplemental References

S1. Andréfouët, S., Muller-karger, F. E., Robinson, J. A., Christine, J., Torres-pulliza, D., Spraggins, S. A., and Murch, B. (2006). Global assessment of modern coral reef extent and diversity for regional science and management applications : a view from space. Coral Reefs *1745*, 1732–1745.

S2. Labrosse, P., Kulbicki, M. and Ferraris, J. (2002). Underwater visual fish census surveys. Proper use and implementation. Reef Resour. Assess. tools. Noumea, New Caledonia Secr. Pacific Community, 54 pp.

S3. Findley, J. S., and Findley, M. T. (2001). Global , regional , and local patterns in species richness and abundance of butterflyfishes. Ecol. Mon *71*, 69–91.

S4. Clua, E., and Legendre, P. (2008). Shifting dominance among Scarid species on reefs representing. Aquat. Living Resour. *348*, 339–348.

S5. Bellwood, D. R., Hoey, A. S., and Hughes, T. P. (2012). Human activity selectively impacts the ecosystem roles of parrotfishes on coral reefs. Proc. Biol. Sci. *279*, 1621–9.

S6. Kulbicki, M., Bozec, Y., Labrosse, P., Letourneur, Y., Mou-tham, G., and Wantiez, L. (2005). Diet composition of carnivorous fishes from coral reef lagoons of New Caledonia. Aquat. Living Resour. *18*, 231–250.

S7. Chabanet, P., Guillemot, N., Kulbicki, M., Vigliola, L., and Sarramegna, S. (2010). Baseline study of the spatio-temporal patterns of reef fish assemblages prior to a major mining project in New Caledonia (South Pacific). Mar. Pollut. Bull. *61*, 598–611.

S8. Kulbicki, M., Cornuet, N., Vigliola, L., Wantiez, L., Moutham, G., and Chabanet, P. (2010). Counting coral reef fishes: Interaction between fish life-history traits and transect design. J. Exp. Mar. Bio. Ecol. *387*, 15–23.

S9. Kulbicki, M., Mou-tham, G., Vigliola, L., Wantiez, L., Manaldo, E., Labrosse, P., and Letourneur, Y. (2011). Major coral reef fish species of the South pacific with basic information on their biology and ecology. Cris. report. Noumea SPC. 107 pp. + Annex.

S10. O'Gorman, E. J., Yearsley, J. M., Crowe, T. P., Emmerson, M. C., Jacob, U., and Petchey, O. L. (2011). Loss of functionally unique species may gradually undermine ecosystems. Proc. Biol. Sci. *278*, 1886–93.

S11. Mouchet, M., and Mason, N. W. H. (2008). Towards a consensus for calculating dendrogram-based functional diversity indices. Oikos *117*, 794 – 800.

S12. Cowman, P. F., and Bellwood, D. R. (2011). Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. J. Evol. Biol. *24*, 2543–62.

S13. Allen, B., Kon, M., and Bar-Yam, Y. (2009). A new phylogenetic diversity measure generalizing the shannon index and its application to phyllostomid bats. Am. Nat. *174*, 236–243.

S14. Clua, E., Legendre, P., Vigliola, L., Magron, F., Kulbicki, M., Sarramegna, S., Labrosse, P., and Galzin, R. (2006). Medium scale approach (MSA) for improved assessment of coral reef fish habitat. J. Exp. Mar. Bio. Ecol. *333*, 219–230.

S15. Sandin, S. A., Vermeij, M. J. A., and Hurlbert, A. H. (2008). Island biogeography of Caribbean coral reef fish. Glob. Ecol. Biogeogr. *17*, 770–777.

S16. Mellin, C., Bradshaw, C. J. A., Meekan, M. G., and Caley, M. J. (2010). Environmental and spatial predictors of species richness and abundance in coral reef fishes. Glob. Ecol. Biogeogr. *19*, 212–222.

S17. Treml, E. A., Halpin, P. N., Urban, D. L., and Pratson, L. F. (2007). Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. Landsc. Ecol. *23*, 19–36.

S18. Kool, J. T., Paris, C. B., Barber, P. H., and Cowen, R. K. (2011). Connectivity and the development of population genetic structure in Indo-West Pacific coral reef communities. Glob. Ecol. Biogeogr. *20*, 695–706.

S19. Leis, J. M., Herwerden, L. V. A. N., and Patterson, H. M. (2011). Estimating connectivity in marine fish populations: what works best? Oceanogr. Mar. Biol. An Annu. Rev. *49*, 193–234.

S20. Bellwood, D. R., and Hughes, T. P. (2001). Regional-scale assembly rules and biodiversity of coral reefs. Science *292*, 1532–5.

S21. Bell, J. D., Kronen, M., Vunisea, A., Nash, W. J., Keeble, G., Demmke, A., Pontifex, S., and Andréfouët, S. (2009). Planning the use of fish for food security in the Pacific. Mar. Policy *33*, 64–76.

S22. Klugman, J., Rodríguez, F., and Choi, H.-J. (2011). The HDI 2010: new controversies, old critiques. J. Econ. Inequal. *9*, 249–288.

23. Knudby, A., Brenning, A., and LeDrew, E. (2010). New approaches to modelling fish–habitat relationships. Ecol. Modell. *221*, 503–511.

S24. Chatfield, B. S., Van Niel, K. P., Kendrick, G. A., and Harvey, E. S. (2010). Combining environmental gradients to explain and predict the structure of demersal fish distributions. J. Biogeogr. *37*, 593–605.

S25. Elmqvist, T., Folke, C., Nyström, M., Peterson, G., Bengtsson, J., Walker, B., and Norberg, J. (2003). Response diversity, ecosystem change, and resilience. Front. Ecol. Environ. *1*, 488–494.

S26. Folke, C., Carpenter, S., Walker, B., Scheffer, M., Elmqvist, T., Gunderson, L., and Holling, C. S. (2011). Regime Shifts, Resilience, and Biodiversity in Ecosystem Management. Rev. Lit. Arts Am. *35*, 557–581.

S27. Graham, M. H. (2003). Confronting multicollinearity in ecological multiple regression. Ecology *84*, 2809–2815.

S28. Elith, J., Leathwick, J. R., and Hastie, T. (2008). A working guide to boosted regression trees. J. Anim. Ecol. *77*, 802–813.

S29. Davies, R. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative - linear model case. Biometrika *74*, 33 – 43.

S30. Muggeo, V. M. R. (2003). Estimating regression models with unknown break-points. Stat. Med. *22*, 3055–71.